Eye in the Sky - Satellite Image Segmentation

Inter-IIT Tech Meet
IIT Kanpur

Apurv Gupta
Harish Rajagopal
Harsh Sinha

December 19^{th} 2018

Outline

Outline

Satellite Imagery

- Developments in satellite imagery presents a challenging task of making sense of this surplus data and draw meaningful, useful inferences.
- One of the major tasks is to classify objects present in the satellite images.
- Deep Learning architectures have proven to beat classical classification algorithms, and humans too, at supervised classification tasks, consequently are now widely used for classification in satellite images.

Outline

Motivation – why analyse Satellite Images?

- Satellite Imagery is being used extensively by governments, researchers and businesses across the globe for surveillance and analysis.
- This information could either be used by cartographers to create custom maps, or by a countrys defence forces for better surveillance.
- It is also used by geologists to monitor several natural phenomena such as oil spills, change in geometries of icebergs and glaciers etc.

Outline

Problem Statement

- The task is to classify all pixels in a given satellite image, as belonging to one of given 9 classes.
- The classes include: Roads, Buildings, Trees, Grass, Bare Soil, Water, Railways, Swimming pools and Unidentified class.
- Each class is represented by a colour to make it visually distinguishable from the other.
- Some pixels in the ground truth are assigned the colour white (unidentified class), which means that the pixel wasnt assigned any label.

Outline

Multi-Label Classification

- Multi-label classification deals with the problem of classifying each data-point in a set as belonging to one of many classes, or clusters.
- The given task naturally fits into a multi-label classification problem, with each of the given 9 classes of objects representing a label.
- We will attempt to solve it by using deep neural networks that systematically extract simple and complex features from the provided images in order to perform classification.

Outline

Semantic Segmentation

- The aim of Image segmentation is to divide the image into homogeneous regions, which are analyzed to locate objects and boundaries (lines, curves, etc) in images.

Semantic Segmentation

- More precisely, is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics.
- We treat Satellite image classification as a two-step approach which involves Semantic Segmentation of images followed by a pixel-level Classification into predefined classes.
- Convolutional Neural Networks (CNNs) are the most widely used architectures for Image segmentation. They are able to learn appropriate feature representation in an end to end manner.

Outline

Data pre-processing

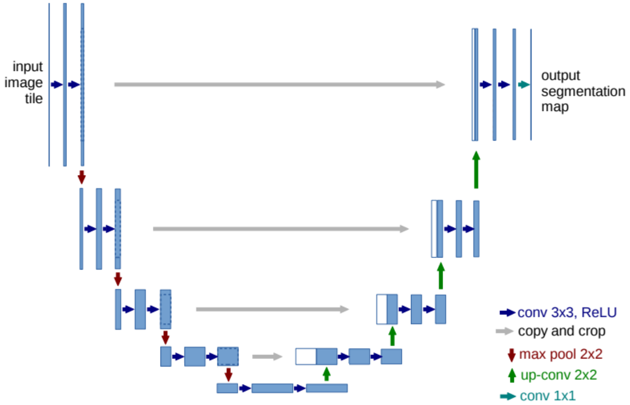
- We are given a training dataset of 14 high quality *.tif satellite images, with corresponding ground truth segmentation maps.
- We tile each image into smaller images. Each tile is then used as a separate image, on which the network is trained.
- We rotate each of the tiles by multiples of 90 degrees in order to augment the dataset.

Validation Procedure

- Since CNNs do not depend on input image size, we can simply use the model trained on tiles to predict for a bigger image.
- In order to quantify the networks capability to generalize on different images, we chose 3 images from the set for testing purposes and trained the network on only 11 images.
- Our metrics for gauging accuracy are: kappa coefficient, confusion matrix, per pixel classification accuracy.
- Note that these metrics are reported only on the unseen 3 images, and are thus good measures of accuracy.

Outline

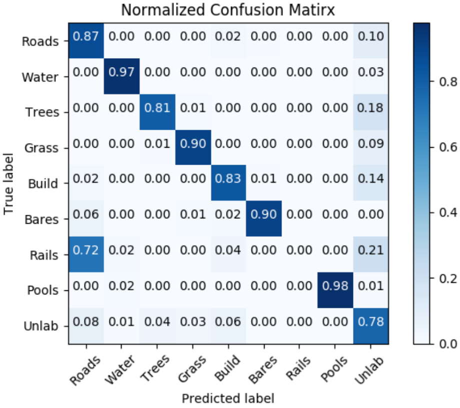
Basic U-Net model: Architecture



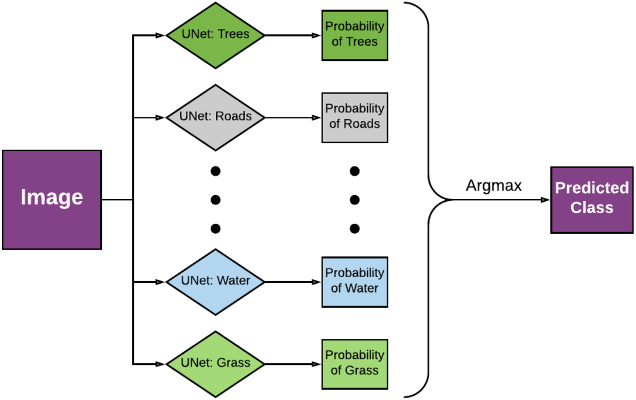
Standard UNet Model as designed by Ronneberger et al. in [1]

Basic U-Net model: Results

Accuracy	76%
Cohen's Kappa	0.7646



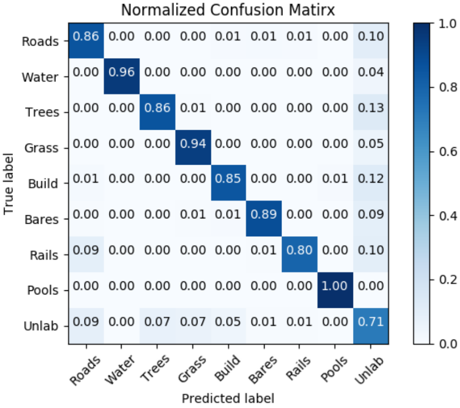
One Vs All U-Net model: Architecture



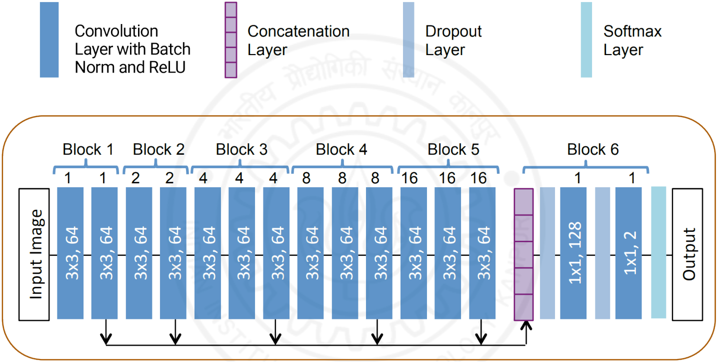
Each individual UNet Block has same architecture as before.

One Vs All U-Net model: Results

Accuracy	91.41%
Cohen's Kappa	0.7577



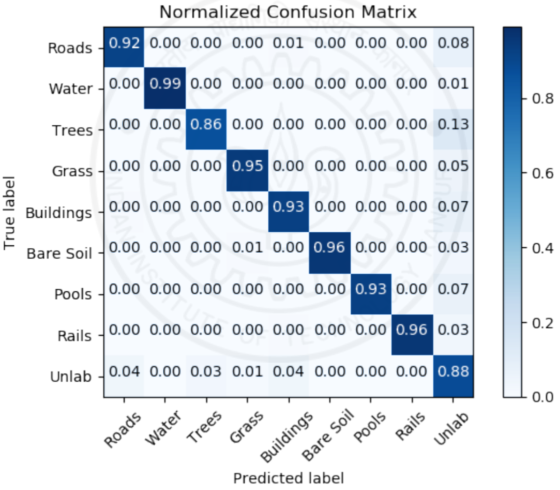
P-Net model: Architecture [2]



The numbers in dark blue boxes denote convolution kernel sizes and numbers of output channels.
The numbers on the top of these boxes denote dilation parameters.
Model is taken from work by Wang et al. [2]

P-Net model: Results

Accuracy	90.68%
Cohen's Kappa	0.8776



Hyperparameter Optimization

We perform hyperparameter search using Tree-Structured Parzen Estimators method [3] using the open source library hyperopt [4]. The next set of hyperparameters are chosen based on the Expected Improvement:

$$EI_{y^*}(x) = \int_{-\infty}^{y^*} (y^* - y)p(y|x)dy$$

where y is the loss and y^* is a threshold value for the loss.

$$\gamma = p(y < y^*), \qquad p(x|y) = \begin{cases} l(x) & \text{if } y < y^* \\ g(x) & \text{if } y \geq y^* \end{cases}$$

where l and g are hierarchical processes. By applying Bayes' Rule and some simplification we get:

$$EI_{y^*}(x) \approx (\gamma + \frac{g(x)}{l(x)}(1 - \gamma))^{-1}$$

Thus to maximise the EI we maximise ratio $\frac{l(x)}{g(x)}$.

Conclusion

- In comparison to the simple Unet the multi-Unet performs much better due to the increase in the number of parameters learnt.
- P-Net and Multi-Unet perform at par in terms of classification accuracy.

Further Improvements

- Generative models such as GANs can be used to synthesise semantically similar data. This solves the problem of paucity of satellite image data. The classification can also be done using the discriminator itself.
- Rotations other than with multiples of 90 degrees can be used for further dataset augmentation. The process would be computationally very expensive but the augmented dataset would be magnitudes of times larger and more effective.
- Post processing on the images produced on the networks would further increase the accuracy for segmentation by removing errors such as incorrect classification of isolated pixels.

References

- Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.
- Wang, Guotai, et al. "Interactive medical image segmentation using deep learning with image-specific fine-tuning." IEEE Transactions on Medical Imaging (2018).
- Bergstra, James S., et al. "Algorithms for hyper-parameter optimization." Advances in neural information processing systems. 2011.
- James Bergstra, Daniel Yamins, and David Daniel Cox. "Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures." 2013
- Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.